

ATTACHMENT 2 to WQB Meeting 9/23/15

Comments and responses for proposed amendments to R317-2 published in the June 1, 2015 Utah Bulletin No. 39397. Only written comments were received. No comments were received at the Public Hearing July 6, 2015.

- 1. Comment:** A comment raises a concern regarding the proposed revision to R317-2-7.1 regarding not listing lakes and streams as impaired when the water quality exceedance was due to background conditions. USEPA Region 8 (USEPA8) notes that the language does not limit the determination that a criterion exceedance is due to background to assessments but would also be applicable to other programs such as permitting.

DWQ Response: DWQ agrees with the comment and is proposing to withdraw the provision as shown below to develop wording that will limit the implementation of this provision to assessments.

~~R317-2-7.....Site-specific standards may be adopted by rulemaking where biomonitoring data, bioassays, or other scientific analyses indicate that the statewide criterion is over or under protective of the designated uses or where natural or unalterable conditions or other factors as defined in 40 CFR 131.10(g) prevent the attainment of the statewide criteria as prescribed in Subsections R317-2-7.2, and R317-2-7.3, and Section R317-2-14. When it is determined that natural background level of a pollutant is less stringent than the otherwise applicable criterion, the water quality criterion will be equal to the natural background concentration.~~

- 2. Comment:** A comment raises a concern regarding the proposed revision to R317-2-7.1 regarding not listing lakes and streams as impaired when the water quality exceedance was due to background conditions. USEPA specifically identifies 2 components (out of 3) that are recommended in USEPA guidance as missing in Utah's standards:
 1. A definition of natural background that only includes non-anthropogenic sources;
 2. A procedure for determining natural background, or alternatively, a reference in their WQS to another document describing the binding procedure that will be used.

DWQ Response: DWQ agrees that a definition for natural background is necessary. DWQ proposes to withdraw the revisions to R317-2-7 regarding background conditions at this time. DWQ will develop a definition for background conditions with the Water Quality Standards Workgroup. With regards to the binding procedures, further discussion with USEPA8 and the Water Quality Standards Workgroup are necessary because of the lack of similar USEPA guidance or specific regulations.

- 3. Comment:** *"We continue to have significant concerns with the methods used to derive the maximum criterion when R317-2-7.1 allows for a 10% exceedance of maximum TDS criteria when making assessment decisions. The Division's approach to deriving site-*

specific maximum criteria is to evaluate a wide range of upper percentile values that are intended to approximate the maximum. The Division has set maximum criteria to three different upper limit statistics. The criteria adopted in 2014 (Blue Creek and Blue Creek Reservoir) were set to the 95 % upper prediction limits (UPL95) for the next 5 observations. The revised seasonal maximum criteria for Blue Creek are set to the USL95 and UTL95-99 for summer and winter, respectively. The ProUCL 5.0 Technical Guide provides the following descriptions of these statistics...

...These statistics either provide high confidence that future samples will be less than the limit (i.e., UPLs and UTL – both with a low false positive rate) or are statistics that are typically used to estimate the true maximum of a given distribution (i.e., USL). Figure 7 from the Division’s support document clearly shows that the proposed maximum criteria are greater than what has been observed in Blue Creek over the last 20+ years. We question why the Division is interested in setting the criterion to an estimate of the true maximum, rather than a percentile of the distribution? Estimating a true maximum is a challenging task that inflates the limit and results in less protective criteria. It is also worthy to note the statistical outlier in the dataset (7,180 mg/L, not presented in these figures) is less than the proposed maximum criteria. Use of the proposed maximum criteria to establish permit limit or when making assessment decisions, which allows for a 10% exceedance, will not protect the existing water quality conditions in Blue Creek.”

DWQ Response 3a. USEPA8 comments that: *“Use of the proposed maximum criteria to establish permit limit or when making assessment decisions, which allows for a 10% exceedance, will not protect the existing water quality conditions in Blue Creek.”*

DWQ disagrees with this conclusion. The Blue Creek TDS seasonal criteria include both an average and a maximum which provides much more rigorous protection than just the maximums currently applied everywhere else in Utah for TDS criteria. The promulgation of the average criteria alone are sufficient to address all of the concerns regarding protectiveness in USEPA8’s comments because hypothetically even if the proposed maximum criteria were too high, the average criteria would protect the water quality of Blue Creek.

However, USEPA8 comments indicate that disapproval would be likely, so DWQ evaluated the impacts of lowering the maximum criteria to decrease the probability of a USEPA8 disapproval. The potential impacts of lower maximum criteria to permits as well as water quality assessments were evaluated and DWQ concluded that impacts of reducing the maximum criteria are acceptable.

As shown in the following text, the maximum seasonal criteria for both Blue Creek and Blue Creek Reservoir were reduced by basing them on the 95% upper tolerance limits of the 90th percentiles. For Blue Creek Reservoir, the revised maximum criterion is 2,100 mg/l (previously 2,200 mg/l). The revised maximum criteria for Blue Creek are 4,900 mg/l (previously 7,200 mg/l) for the summer season summer and 6,700 mg/l (previously 7,500 mg/l) for the winter season.

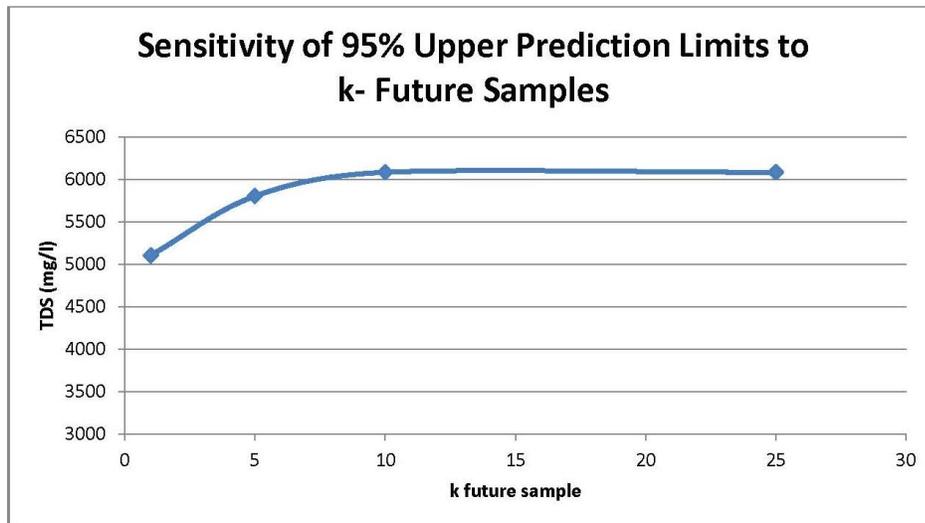
For permits, the average criteria are expected to remain the primary criteria for protecting water quality. The maximum criteria are more stringent than the previously proposed maximums but based on discussions with the permittee, the impacts are expected to be minimal.

For assessments, the primary concerns with the reduced maximum criteria are false-positive water quality impairments. In simulations, additional sampling, although an undesirable expenditure of resources, appears to be sufficient to address false positive water quality impairments if they occur in the future.

Upper prediction limits were also considered but rejected for the revised maximum criteria. Staff evaluated the effects of the k-number of future comparison samples (e.g., samples collected for assessment) on the upper prediction limits. As shown on the following figure, the prediction limits change depending on the number of future samples. Based on these observations, upper prediction limits were not considered optimal for the maximum criteria. For the same reason, the upper prediction limits included in the average criteria for assessment were deleted. The *Site-Specific Standard for Total Dissolved Solids Blue Creek Reservoir and Blue Creek, September 3, 2015* (2015 criteria support document) was revised to provide information and data for statistically rigorous future assessments of the average criteria. This approach will allow assessment methods to be optimized based on the number of samples of available. The assessment methods will be documented in *Integrated Reports*.

Blue Creek and tributaries, Box Elder County, from ~~Gunnison~~ Bear River Bay, Great Salt Lake to Blue Creek Reservoir: maximum ~~6,300 mg/l~~ and an ~~average of 3,900 mg/l~~, March through October daily maximum 4,900 7,200 mg/l and an average of 3,800 mg/l; November through February daily maximum 6,300 7,500 mg/l and an average of 4,700 mg/l. Assessments will be based on TDS concentrations measured at the location of STORET 4960740. ~~At least 10 samples are required to assess compliance with the average criterion. If the sample average for samples collected from March through October is equal to or less than 4,100 mg/l and the sample average for samples collected from November through February is equal to or less than 5,300 mg/l, the average criteria are met. Alternative scientifically defensible assessment methods may be applied for assessing the average criteria.~~

Blue Creek Reservoir and tributaries, Box Elder County,
daily maximum 2,100 2,200 mg/l



DWQ Response 3b. EPA comments that *“These statistics either provide high confidence that future samples will be less than the limit (i.e., UPLs and UTL – both with a low false positive rate) or are statistics that are typically used to estimate the true maximum of a given distribution (i.e., USL).”* And *“We question why the Division is interested in setting the criterion to an estimate of the true maximum, rather than a percentile of the distribution? Figure 7 from the Division’s support document clearly shows that the proposed maximum criteria are greater than what has been observed in Blue Creek over the last 20+ years.”*

USEPA8 appears to have misinterpreted the statistical parameters proposed. The ProUCL statistical program was developed to be used primarily by the Superfund Program for cleaning up abandoned hazardous waste sites. At these sites, background concentrations of highly toxic metals and metalloids (for instance, arsenic) often need to be characterized to limit any clean up to contamination. In these situations, false negative decision errors (concluding that an area is not contaminated when it is) are highly undesirable and the methods recommended by ProUCL balance the potential for both false positive and negative decisions. This is apparent from the following statements in the USEPA (2013) ProUCL guidance:

“Caution: To provide a proper balance between false positives and false negatives, the upper limits described above, especially a 95% USL (USL95) should be used only when the background data set represents a single environmental population without outliers (observations not belonging to background). Inclusion of multiple populations and/or outliers tends to yield elevated values of USLs (and also of UPLs and UTLs) which can result in a high number (and not necessarily high percentage) of undesirable false negatives, especially for data sets of larger sizes (e.g., $n > 30$).”[p. 86]

“Notes: The user specifies the allowable false positive error rate, $\alpha (=1-CC)$, and the false negative error rate (declaring a location clean when in fact it is contaminated) is controlled by making sure that one is dealing with a defensible/established background data set representing a single background population and the data set is free of outliers.”[p. 87]

As documented in the revised 2015 criteria support document, the analyses for Blue Creek adhered to these recommendations. DWQ acknowledges that the maximum, or any criterion, are more protective the lower they are set. The minimum requirement for this site-specific standard is to be as protective of the natural conditions as modified by irreversible conditions. By meeting this requirement, the criteria will not allow water quality to be degraded and hence will be protective. Setting the criteria more stringent will provide additional protection but also may result in undesirable outcomes such as false-positive impairment decisions that result in resources being unnecessarily diverted to address the “impairment”. Setting the criteria too low could also result in unnecessarily stringent permits resulting in permittees incurring unnecessary treatment costs. Therefore, and as documented in the 2015 criteria support document, DWQ has applied methods that balance both false positive and negative decision errors.

All of the potential parameters considered by DWQ were percentile estimates and no estimates of the maximum were proposed by DWQ, nor are methods for estimating the maximum provided in the ProUCL guidance (USEPA, 2013). The 2015 criteria support document does indicate that the true maximum would be appropriate for the maximum criterion. The 2015 criteria support document was revised to indicate that estimates of the corresponding duration and frequency are desirable parameters along with the true maximum. Because a maximum could not be determined, DWQ proposed statistical estimates of the 95th percentile or higher. USEPA8 appears to believe that because the proposed maximum criteria are higher than any of the existing observations that the proposed maximums are too high. DWQ disagrees because it is statistically improbable that the maximum was sampled and is therefore included in the existing data.

Consider the following simplified example. There would only be a 4% chance that a sample was collected on the day with the maximum TDS concentration if samples are collected on 349 days out of 8,740 possible days (sampled days and possible sample days for Blue Creek data). Therefore, it is highly unlikely that the sample data set includes the daily true maximum.

4. **Comment:** “To resolve our concerns with the proposed approach, we suggest that the Division consider adding an additional statement to the site-specific standard that when making assessment decisions, the 10% exceedance frequency in R317-2-7.1 does not apply to the maximum criteria. This approach would only address the concerns with assessment decisions and does not address the implementation in UPDES permits; however, it is likely that permit limits derived using the average criteria will control effluent concentrations such that the maximum criterion will never be observed. Alternatively, the Division could consider an approach similar to what is proposed for the average criteria where the statistical uncertainty with the dataset is taken into consideration in the assessment thresholds, rather than the water quality criterion. The UPL/UTL/USL limits are more akin to the assessment thresholds than values that are expected to protect the existing water quality of Blue Creek. The maximum criterion could then be set to a more protective limit that is compatible with R317-2-7.1 (e.g., 90th percentile or potentially maximum observed, depending on the robustness of the dataset).”

DWQ Response. DWQ has elected not to exclude the Blue Creek site-specific standards from the 10% allowance in R317-2-1 because this provision is clearly intended to apply to site-specific TDS standards:

“For water quality assessment purposes, up to 10 percent of the representative samples may exceed the minimum or maximum criteria for dissolved oxygen, pH, E. coli, total dissolved solids, and temperature, including situations where such criteria have been adopted on a site-specific basis.”

Instead, DWQ has set the maximum criteria at the 90th percentile (95% upper tolerance limit of the 90th percentile). For the reasons discussed in response to comment 3, the upper prediction limit assessment thresholds have been deleted from the average criteria for Blue Creek. Specific assessment methods will be consistent with how the average criteria were derived and based on the assessment sampling design.